E1 244: Detection & Estimation Theory **Jan-May 2017**

Lecture 4 — January 12

Lecturer: Aditya Gopalan *Scribe: Raseena KT, Shanuja Sasi*

Continuation of Minimax Hypothesis Testing

4.1 Recap

We assume that there are two possible hypotheses, H_0 and H_1, which correspond to two probability distributions $\mathbb{P}_0$ and $\mathbb{P}_1$, respectively, on $(\Gamma, \mathcal{G})$ where Γ is the observation set and $\mathcal{G}$ is the class of subsets of Γ that can be assigned probabilities.

4.2 Minimax Hypothesis Testing

The structure of minimax rules has already been stated in the lecture 3.
For case (3), i.e, for $\mathbb{R}_0(\delta_{\pi_L}) = \mathbb{R}_1(\delta_{\pi_L})$, suppose v(π_0) is differentiable at its maximum π_L.

$$\begin{aligned} V(\pi_0) &\leq r(\pi_0, \delta_{\pi_L}) \quad \forall \;\; \pi_0 \in [0,1], \\ V(\pi_L) &= r(\pi_L, \delta_{\pi_L}). \end{aligned}$$

Therefore the line r(. ,δ_{π_L}) is tangent to V(π_0) at π_L.If V(π_0) is differentiable at $\pi_L \in$ (0,1).

$$\begin{aligned} 0 &= V'(\pi_L) \\ &= \frac{d}{d\pi_0} V(\pi_0) \,|_{\pi_0=\pi_L} \\ &= \frac{d}{d\pi_0} r(\pi_0 \delta_{\pi_L}) \,|_{\pi_0=\pi_L} \\ &= R_0(\delta_{\pi_L}) - R_1(\delta_{\pi_L}) \end{aligned}$$

which implies

$$R_0(\delta_{\pi_L}) = R_1(\delta_{\pi_L}).$$

In all the above Minimax Hypothesis testing approach we have assumed that the function $V(\pi_0)$ is continuous at the point of maxima.This may not be the case in many problems of interest. From the below tree structure , it can be seen that minimax rules are designed for all the cases except when the function V is not differentiable at its maximum interior point π_L.

π_L:Maximizer of V

$\pi_L \in \{0, 1\}$

δ_{π_L} is the minimax rule

$\pi_L \in (0, 1)$

V is differentiable at π_L

δ_{π_L} is the minimax rule

V is not differentiable at π_L

?

Example 4.2.1. Consider the binary channel hypothesis testing discussed in lecture 2. The observation space Γ contains only two elements 0 and 1. Therefore there are only four possible decision rules namely $\delta_1, \delta_2, \delta_3$ and δ_4. It is illustrated in the figure 4.2. It can be seen that the function V is not differentiable at its maxima π_L. In cases like this a randomized decision rule δ^* solves the minimax problem.

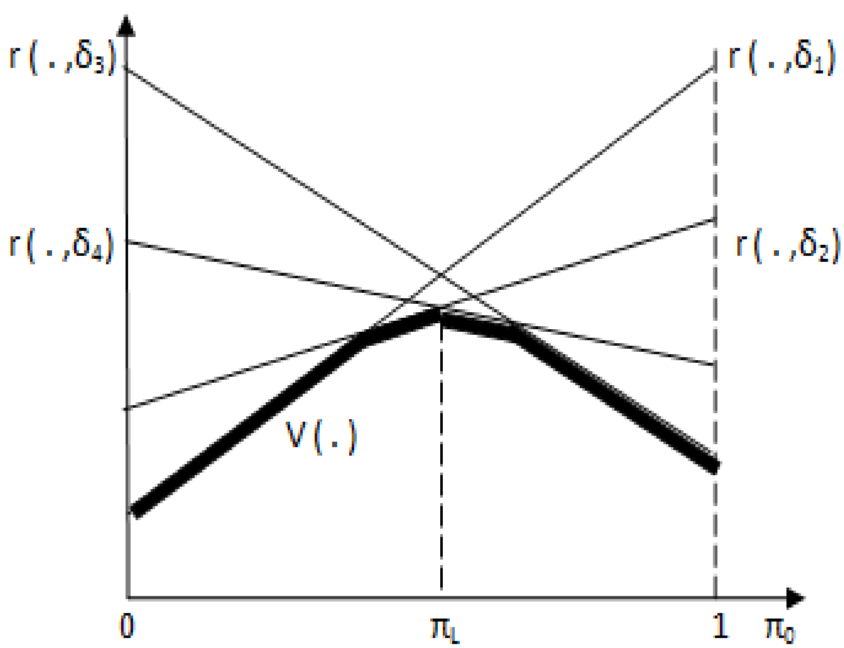

Figure 4.1: Binary channel hypothesis testing

Definition 4.2.2. Randomized decision rule δ^* maps $\Gamma \to [0,1]$.

$$\delta^* : \Gamma \to [0,1]$$

Interpretation:After having seen y $\in \Gamma$, the rule decides H_1 with probability δ(y) and H_0 with probability 1 - δ(y).

Assumption 4.2.3. Assume $C_{11} - C_{01} < 0$ and $C_{00} - C_{10} < 0$ without loss of generality.

For any $\pi_0 < \pi_L$, the bayes rule is

$$\Gamma_1(\delta_{\pi_0}) = \{y \in \Gamma : L(y) \geq \tau(\pi_0)\},$$

where

$$\tau(\pi_0) = \frac{\pi_0(C_{00} - C_{10})}{(1-\pi_0)(C_{11} - C_{01})}$$

$\tau(\pi_0)$ is a monotonically increasing function of π_0. And hence $\forall\ \pi_0' < \pi_0'' < \pi_L$, we have

$$\Gamma_1(\delta_{\pi_0'}) \subseteq \Gamma_1(\delta_{\pi_0''})$$

$$\begin{aligned}
\text{Define:} \qquad \Gamma_1^- &= \bigcap_{\pi_0 < \pi_L} \{\Gamma_1(\delta_{\pi_0'})\} \\
&= \bigcap_{\pi_0 < \pi_L} \{L(y) \geq \tau(\pi_0)\} \\
&= \{y \in \Gamma : L(y) \geq \tau(\pi_L)\}. && (4.0) \\
\Gamma_1^+ &= \bigcup_{\pi_0 < \pi_L} \{\Gamma_1(\delta_{\pi_0'})\} \\
&= \bigcup_{\pi_0 < \pi_L} \{L(y) > \tau(\pi_0)\} \\
&= \{y \in \Gamma : L(y) > \tau(\pi_L)\}. && (4.1)
\end{aligned}$$

The difference in the inequalities of the equations (4.15) and (4.16) is due to the following reason:

$$\bigcap_{\pi_0 < 0} [\pi_0, \infty) = [0, \infty) = \{y : y \geq 0\}, \qquad (4.2)$$

$$\bigcup_{\pi_0 > 0} [\pi_0, \infty) = (0, \infty) = \{y : y < 0\}. \qquad (4.3)$$

Equation (4.17) is similar to (4.15) and (4.18) is similar to (4.16) and hence the difference in the inequalities arises.

From the equations (4.15) and (4.16)

$$\Gamma_1^- \supseteq \Gamma_1^+$$

which implies,

$$\Gamma_1^{-c} \subseteq \Gamma_1^{+c}$$

The observation space is illustrated in the figure 4.3.

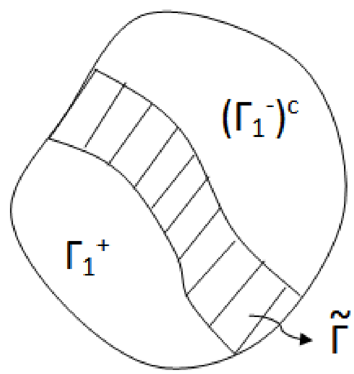

Figure 4.2: Division of the observation space Γ

Take some $q \in [0, 1]$ and define the randomized rule δ_q as:

$$\delta_q(y) = \begin{cases} 1, & \text{if } \; y \in \Gamma_1^+, \\ 0, & \text{if } \; y \in (\Gamma_1^-)^c, \\ \begin{cases} \text{1 with probability q} \\ \text{0 with probability 1-q} \end{cases} & \text{if } \; y \in \widetilde{\Gamma}. \end{cases} \tag{4.4}$$

Let δ^+ be the decision rule that divides observation space Γ into Γ_1^+ and $(\Gamma_1^+)^c$ and δ^- be the decision rule that divides observation space into Γ_1^- and $(\Gamma_1^-)^c$.

Claim 1: Bayes risk of δ_q = bayes risk of δ^+ = bayes risk of δ^-.

Proof. This is because boundary decision is irrelevant to bayes risk. ■

Since conditional risk depends on the boundary condition, it can be written as,

$$R_j(\delta_q) = qR_j(\delta^-) + (1-q)R_j(\delta^+), \qquad j \in \{0, 1\} \tag{4.5}$$

which is nothing but,

$$R_0(\delta_q) = qR_0(\delta^-) + (1-q)R_0(\delta^+),$$
$$R_1(\delta_q) = qR_1(\delta^-) + (1-q)R_1(\delta^+).$$

So for the minimax rule at equality,

$$\max\{R_0(\delta_q), R_1(\delta_q)\} = R_0(\delta_q) = R_1(\delta_q). \tag{4.6}$$

Therefore, the conditional risk equations for equality condition are,

$$\begin{aligned}
&qR_0(\delta^-) + (1-q)R_0(\delta^+) = qR_1(\delta^-) + (1-q)R_1(\delta^+),\\
&q\{R_0(\delta^-) - R_0(\delta^+)\} + R_0(\delta^+) = q\{R_1(\delta^-) - R_1(\delta^+)\} + R_1(\delta^+),\\
&R_0(\delta^+) - R_1(\delta^+) = q\{R_0(\delta^+) - R_1(\delta^+) + R_1(\delta^-) - R_0(\delta^-)\}.
\end{aligned}$$

So,

$$q = \frac{R_0(\delta^+) - R_1(\delta^+)}{R_0(\delta^+) - R_1(\delta^+) + R_1(\delta^-) - R_0(\delta^-)}. \tag{4.7}$$

This gives the probability with which we choose the alternate hypothesis at the boundary, which is equivalent to throwing a biased coin with probability q and selecting a decision rule based on the outcome.

Remark 1. We now represent the unconditional risk by V. Since V is concave, it must have left hand and right hand derivative at π_L, which is denoted by $V'(\pi_L^-)$ and $V'(\pi_L^+)$. Now,

$$\begin{aligned}
V'(\pi_L^-) &= \frac{d}{d\pi_0} r(\pi_0, \delta_{\pi_L}^-)\\
&= \frac{d}{d\pi_0}\{\pi_0 R_0(\delta_{\pi_L}^-) + (1-\pi_0)R_1(\delta_{\pi_L}^-)\}\\
&= R_0(\delta_{\pi_L}^-) - R_1(\delta_{\pi_L}^-).
\end{aligned} \tag{4.8}$$

Similarly, $\quad V'(\pi_L^+) = R_0(\delta_{\pi_L}^+) - R_1(\delta_{\pi_L}^+).$
Hence eqn. (4.7) can be written as

$$q = \frac{V'(\pi_L^+)}{V'(\pi_L^+) - V'(\pi_L^-)}. \tag{4.9}$$

We can analyze the importance of the above equation in the following manner. Figure 4.3 shows the case in which V is discontinuous at the point of maxima. The decision rule here is δ_q. By varying the probability q from 0 to 1, different slopes of the line $r(\pi_0, \delta_{\pi_L})$ can be obtained. The particular value q obtained from the equation above corresponds to the horizontal line.

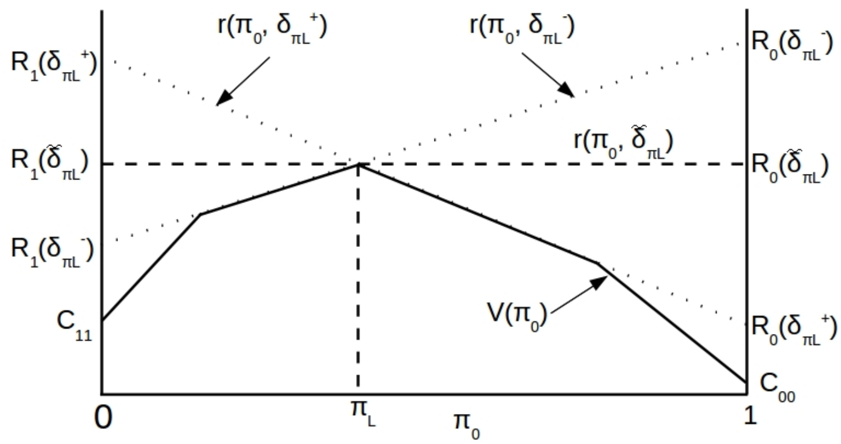

Figure 4.3: Action of the randomized decision rule

Hence we can complete the tree structure as below:

π_L:*Maximizer of V*

$\pi_L \in \{0,1\}$ — δ_{π_L} *is minimax rule*

$\pi_L \in (0,1)$

- *V is differentiable at* π_L — δ_{π_L} *is minimax rule*
- *V is not differentiable at* π_L — δ_q *is the minimax rule*

Example 4.2.4 (Location testing with Gaussian error). The input to the channel is real numbers.Here there are two possible inputs μ_0 and μ_1. The gaussian noise is added to the input and the output is produced. The goal here is to guess which of μ_0 or μ_1 was sent.

$$\Gamma := \Re$$
$$H_j = \mu_j \in \Re \text{ is the input, } j \in \{0,1\}$$
$$\mathbb{P}_j = N(\mu_j, \sigma^2), j \in \{0,1\}$$

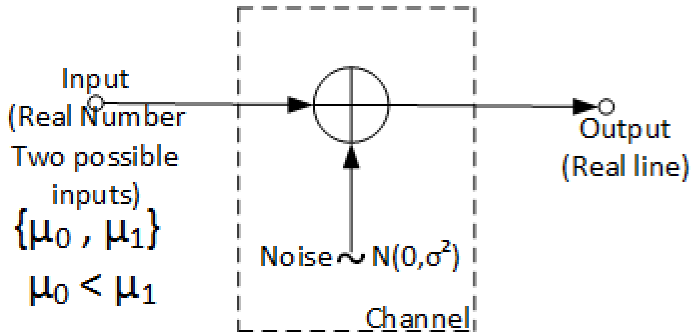

Figure 4.4: Location Testing with Gaussian Error.

Consider the case of uniform costs, i.e, $\mathcal{C}_{ij} = \mathbb{1}\{i \neq j\}$

$$V(\pi_0) = \pi_0 R_0(\delta_{\pi_0}) + \pi_1 R_1(\delta_{\pi_0})$$

δ_{π_0} has the form:

$$\begin{aligned}\Gamma_1 &= \{y \in R : L(y) \geq \tau(\pi_0)\} \\ &= \{y : y \geq \tau'\}.\end{aligned}$$

where L(y)=$\frac{p_1(y)}{p_0(y)}$,
$p_1(y)$:Probability that y is produced by hypothesis 1
$p_0(y)$:Probability that y is produced by hypothesis 0

From the example 2.2.2 from lecture 2,

$$\tau' = \frac{\mu_0 + \mu_1}{2} + \frac{\sigma^2 \ln \tau}{\mu_1 - \mu_0}$$

$$\begin{aligned}V(\pi_0) &= \pi_0 \int_{\tau'}^{\infty} P_0(y)\, dy + (1-\pi_0) \int_{-\infty}^{\tau'} P_1(y)\, dy \\ &= \pi_0[1 - \phi(\frac{\tau' - \mu_0}{\sigma})] + (1-\pi_0)\phi(\frac{\tau' - \mu_1}{\sigma}).\end{aligned} \tag{4.10}$$

ϕ is the CDF (Cumulative Distribution Function) of standard normal random variable.
It turns out that V(π_0) is maximized at π_0= $\frac{1}{2}$.So the minimax rule is simply $\delta_{\frac{1}{2}}$. i.e,

$$\delta_{\frac{1}{2}} = \mathbb{1}\{y \geq \frac{\mu_0 + \mu_1}{2}\}.$$